



ELSEVIER

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp



Rule-based category use in preschool children



Fabien Mathy^{a,*}, Ori Friedman^b, Brigitte Courenq^a, Lucie Laurent^{a,c},
Jean-Louis Millot^d

^a BCL Lab, UMR 7320, Department of Psychology, Université Nice Sophia Antipolis, 06357 Nice, France

^b Department of Psychology, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

^c Maison des Sciences de l'Homme et de l'Environnement Ledoux, Université de Franche-Comté, 25030 Besançon cedex, France

^d Neurosciences Laboratory of Besançon EA-481, Université de Franche-Comté, 25030 Besançon cedex, France

ARTICLE INFO

Article history:

Received 14 February 2014

Revised 16 October 2014

Keywords:

Rule use

Mental flexibility

Rules

Categorization

Development

Boolean complexity

ABSTRACT

We report two experiments suggesting that development of rule use in children can be predicted by applying metrics of complexity from studies of rule-based category learning in adults. In Experiment 1, 124 3- to 5-year-olds completed three new rule-use tasks. The tasks featured similar instructions but varied in the complexity of the rule structures that could be abstracted from the instructions. This measure of complexity predicted children's difficulty with the tasks. Children also completed a version of the Advanced Dimensional Change Card Sorting task. Although this task featured quite different instructions from those in our "complex" task, performance on these two tasks was correlated, as predicted by the rule-based category approach. Experiment 2 predicted findings of the relative difficulty of the three new tasks in 36 5-year-olds and also showed that response times varied with rule structure complexity. Together, these findings suggest that children's rule use depends on processes also involved in rule-based category learning. The findings likewise suggest that the development of rule use during childhood is protracted, and the findings bolster claims that some of children's difficulty in rule use stems from limits in their ability to represent complex rule structures.

© 2014 Elsevier Inc. All rights reserved.

* Corresponding author.

E-mail address: fabien.mathy@unice.fr (F. Mathy).

Introduction

Many rules guide people's behavior in everyday life. These include rules of courtesy and politeness (e.g., when asking for something, say "please"), of law (e.g., stop on red, go on green), of safety (e.g., do not touch a hot stove), and of games and sports (e.g., collect \$200 when you pass "go"). Rule systems often change across contexts, so people must often override or disregard previously relevant rules and flexibly adopt current ones.

Children often have difficulty in using simple rules, and the ability to successfully follow rules improves with age. For example, preschool-aged children have difficulty with a variety of tasks in which using rules requires avoiding dominant response tendencies (Zelazo & Carlson, 2012). This can be observed in the day–night task, where children have difficulty in following rules requiring them to say "night" to pictures showing the sun and "day" to pictures showing the moon (Diamond, Kirkham, & Amso, 2002; Gerstadt, Hong, & Diamond, 1994); this difficulty presumably arises because these rules conflict with children's more dominant tendencies to say "day" for the sun and "moon" for the night. Another example is children's performance on the Dimensional Change Card Sort (DCCS) task. In this task, children first follow one sorting rule to sort cards according to one of two dimensions (e.g., color) but then switch rules and sort according to a second dimension (e.g., shape). Whereas 3-year-olds mostly fail to make this switch and continue sorting using the first rule, 4- and 5-year-olds succeed in switching (Frye, Zelazo, & Palfai, 1995; Hanania & Smith, 2009; van Bers, Visser, van Schijndel, Mandell, & Raijmakers, 2011; Zelazo, Frye, & Rapus, 1996; Zelazo, Müller, Frye, & Marcovitch, 2003a). However, difficulties remain for these older children in the advanced version of the task, which requires switching between the shape and color rules on consecutive trials (Carlson, 2005; Chevalier & Blaye, 2009; Hongwanishkul, Happaney, Lee, & Zelazo, 2005). Similar difficulties even arise for adults if we consider their response times (Diamond & Kirkham, 2005).

Rule-based category learning

In the current article, we suggest that insight into the development of rule use in children can be gained from an existing literature on rule-based category learning. This field has mostly sought to explain adults' difficulties in learning various artificial rule-based categories (for the seminal studies, see Shepard, Hovland, & Jenkins, 1961, and Medin & Schaffer, 1978; for more recent important developments, see Nosofsky, Gluck, Palmeri, McKinley, & Gauthier, 1994, and Rehder & Hoffman, 2005). Some articles have examined rule-based category learning in children as well (e.g., Minda, Desroches, & Church, 2008).

To explain how rule-based categories are learned and represented, the field has developed a formalism based on Boolean complexity minimization. This formalism allows rule-based categories to be represented using logical disjunctive normal formulas such as "*a* and *b* OR *c*." "*Elephant = huge animal with a trunk with large ears (if African) OR small ears (if Indian)*" and "*my favorite pet = white cat OR black dog*" are examples of disjunctive normal forms. A disjunction is a logical formula that expresses categories for which objects do not resemble one another, which automatically increases the complexity of a category (Mathy, Haladjian, Laurent, & Goldstone, 2013). Nearly all studies on rule-based category learning have focused on complex rule-based categories with a minimum of three dimensions. This includes both more recent studies in this field (Bradmetz & Mathy, 2008; Feldman, 2000, 2003b; Lafond, Lacouture, & Mineau, 2007; Minda et al., 2008; Vigo, 2006) and older studies (Bourne, 1970; Bruner, Goodnow, & Austin, 1956; Hovland, 1966; Levine, 1966; Shepard et al., 1961). Because we aim to apply this work to rule use in preschoolers, we instead focus on two-dimensional artificial categories that are used to classify two-dimensional stimuli such as "red square" and "dark flower."

Such formulas are thought to represent the product of an abstraction process. They allow people to build rule-based categories from simple and independent features. Consider a set of four kinds of objects varying only in color and shape: *dark flower*, *light flower*, *dark butterfly*, and *light butterfly*. For this set, one simple rule-based category is "*dark*," which is a minimization of the "*dark flower*, *dark butterfly*" set of objects. This category can be used to classify the four objects into two groups by

considering only one dimension, color (i.e., is the item dark or not?), without considering how this dimension interacts with the other dimension, shape (see Fig. 1 for a depiction of this simple rule structure). Hence, using this category is more efficient than achieving the same classification using a separate categorization rule for each object (i.e., If *dark flower*, then Category A, else Category B; If *dark butterfly*, then Category A, else Category B), a strategy that requires rote memorization and considering both dimensions without any kind of abstraction.

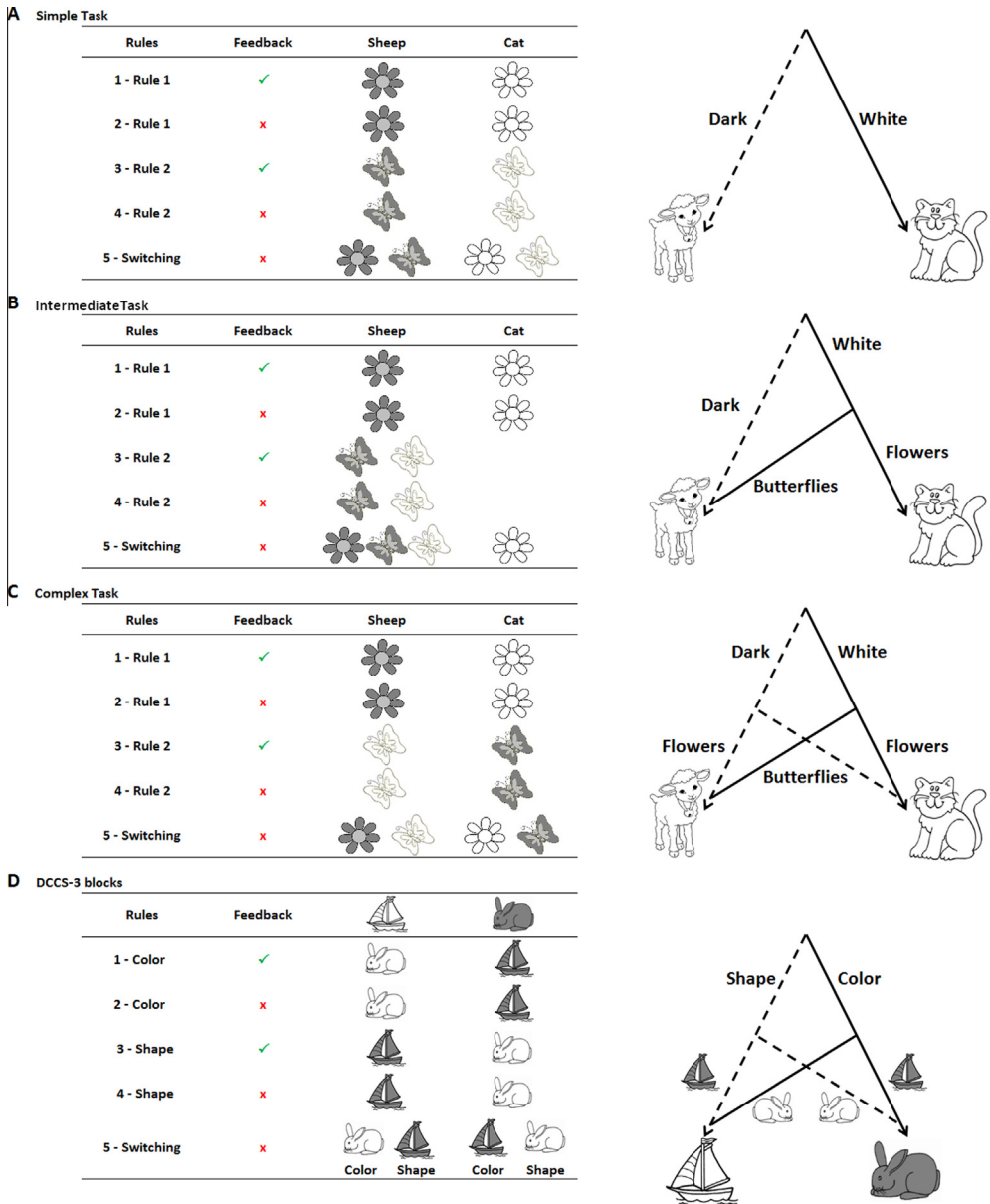


Fig. 1. The three rule-use tasks (Simple, Intermediate, and Complex) and DCCS (Frye, Zelazo, & Palfai, 1995) used in Experiment 1, described by blocks. Rule complexity is represented by decision trees.

A more complex category structure, applying to the same set of objects, is “dark flower OR butterfly” (it can also be represented as “NOT light flower” or as “dark OR butterfly”). This category can be viewed as the minimization of the category set “dark flower, dark butterfly, white butterfly” and allows these objects to be classified into a separate category from objects in the set “light flower.” As can be seen in Fig. 1, using this rule-based category allows half of the objects to be classified by shape alone (butterflies), whereas the other half are classified on both the dimensions of shape and color (flowers). This rule structure involves a “partial interaction” because it requires considering the interaction of both dimensions for half of the cards (Mathy & Bradmetz, 2004). Another way to represent the rule (using the same structure but a different order) is to consider that half of the objects are classified by color alone (dark), whereas the other half are classified on both the dimensions of color and shape (white). In both cases, the white flowers require a two-step process.

One of the most complex two-dimensional categories is “dark flower OR light butterfly.” This category represents the category set “dark flower, light butterfly” and allows these objects to be classified separately from objects in the category set “dark butterfly, white flower.” However, this category is an instance of null minimization because the number of objects it applies to is the same as the number of features in the “minimized” rule. As can be seen in Fig. 1, classifying with this category requires applying reverse rules to flowers and butterflies. Dark flowers go in Category A, and white flowers go in Category B; however, dark butterflies go in Category B, and white butterflies go in Category A. Hence, the category involves a “total interaction” because it requires considering the interaction of both the dimensions of color and shape for all items (Mathy & Bradmetz, 2004). Overall, using two Boolean dimensions, only these three different task structures can be built; here we call these the Simple, Intermediate, and Complex tasks, respectively.

Previous studies on rule-based category learning in both children (ages 4–12 years) and adults (Bradmetz et al., 2008; Mathy, 2012) showed that this complexity metric predicts both learning times across problems and response times across stimuli when the task was to discover the rules by an inductive process. The effect of complexity on response times (which were measured after the rules were correctly learned by the group of adults) is also particularly supportive of the idea that such stimuli cannot simply be categorized using rote memorization, in which case no variance of response times would be observed between different stimuli (i.e., one single step would be sufficient to associate any stimulus with the correct category).

Application to rule use in children

Experiments on rule-based category learning are very different from rule-use experiments (e.g., experiments using tasks like the DCCS and day–night tasks). In rule-use experiments, children are directly told explicit verbal rules and then use them to classify objects or make other responses. In contrast, in typical rule-based category learning experiments, participants are not told explicit verbal rules for sorting. Instead, they attempt to learn these rules based on feedback given after they attempt to sort stimuli.

Despite these differences, we think that the categorization rules (described by rule-based category learning researchers) are useful for study task demands in children’s rule use. Suppose that children are shown stimuli like those discussed above and are told separate rules for sorting each kind of stimulus (e.g., “dark flowers go to Place A”). Children might minimize the rules in representing and applying them. For example, they might minimize the rules “dark flowers go to Place A” and “dark butterflies go to Place A” into the simpler categorization rule “dark go to Place A.” If children do minimize the rules in this way, the complexity metric might predict the difficulty of various rule-use tasks.

Findings consistent with this prediction would be important for several reasons. First, they would suggest that children’s rule use might depend on processes also involved in rule-based category learning. Rule-based categorization has been studied extensively but as an independent topic. Again, a chief difference between these areas is that whereas children are typically told the categorization rules in rule-use studies (including the current experiments), in rule-based category learning tasks participants are instead shown the stimuli and must learn a category representation by induction. As noted above, the rule-based categorization literature has focused on conducting experiments involving at least three dimensions: in children (e.g., Minda et al., 2008), in adults (e.g., Feldman, 2000; Shepard

et al., 1961), and in monkeys (e.g., Smith, Minda, & Washburn, 2004). However, children's use of more basic two-dimensional rule structures has not been studied in preschoolers. Second, such findings would suggest that the development of rule use is protracted during childhood and beyond. For instance, Feldman (2000) reported a catalog of 3 two-dimensional rule-based category structures (those that are studied in the current article) but also 13 three-dimensional and 237 four-dimensional structures, which represent a large playing field for the development of rule use. Third, such findings would bolster claims that at least some of children's difficulty in rule use derives from limits of their ability to represent more complex rule structures (Zelazo & Frye, 1998).

The current approach

The current experiments tested whether the rule-based category approach successfully predicts the difficulty of rule use in young children. Experiment 1 tested this in two ways. First, it compared the relative difficulty of three new rule-use tasks in preschool children aged 3–5 years. In all three tasks, children were told rules for sorting four kinds of bidimensional cards to either of two categories. Specifically, these rules assigned cards to be given to either of two animals: a sheep or a cat (e.g., "The sheep likes the *dark butterflies*"). Because children were told two rules in each task (i.e., a separate rule for each of the two categories), the tasks were matched in complexity at the surface level. However, the tasks varied in the complexity of the categorization rules that could be abstracted from these rules; these are the rules already discussed and depicted in Fig. 1. It was expected that task difficulty would be predicted by this difference between the tasks.

Second, children were also tested on a version of the Advanced DCCS. Although at a surface level this task differs from the other tasks, at a deeper level, it is structurally similar to the most complex of our new tasks. Fig. 1D depicts the rule structure of a version of the DCCS in which participants categorize items by color (e.g., *white rabbits* matched with the *white boat*, *dark boats* matched with the *dark rabbit*) or by shape (e.g., *dark boats* need to be matched with the *white boat*, *white rabbits* need to be matched with the *dark rabbit*).

One can notice in Fig. 1 that the decision trees of our Complex task and the DCCS task are similar in shape. In the DCCS, the first level of the decision tree indicates the game being played (shape or color); the second level indicates the correct categorization rule. In the Complex task, the first level corresponds to the first dimension (color), which enables the stimuli to be categorized according to their shape using the second level (note that the two dimensions can be reversed in the tree; the first level can be associated with shape instead of color without changing the structure of the tree). This suggests that the Complex task and the DCCS share the same decision structure. Hence, we expected these tasks to be of similar difficulty, and we likewise expected performance across these two tasks to be correlated.

Although the Complex task and the Advanced DCCS share the same decision structure, there are important differences between these tasks. First, the tasks differ in the number of stimulus cards used and in the number of rules memorized. The Complex task uses four different kinds of test cards, and children must apply two rules for these cards (i.e., one sorting rule is assigned for each pair of target cards). In contrast, the DCCS uses only two different kinds of test cards, and children must apply two rules for these cards (i.e., both rules apply to both cards). As a consequence of this difference, the Complex task uses different stimuli for the first and second rule-learning phases, whereas the DCCS uses the same stimuli in both rule-learning phases. A second difference is that although both tasks feature materials that factorially vary on two dimensions (i.e., two different shapes that appear in two different colors), this two-dimensionality manifests itself differently in the tasks. In the Complex task, it occurs only in the test cards that children sort and does not apply to the target cards (i.e., the sheep and cat used to indicate where the test cards should be sorted). In this task, the target cards do not share any features with the test cards and are only arbitrarily related to them. In contrast, in the Advanced DCCS, there are only two kinds of target cards, but they do share the features of the test cards, such that children see only a factorial crossing of the two dimensions (shape and color) when looking at test cards in relation to the target cards. These differences between the tasks likely make different demands on children. For instance, the DCCS may make fewer memory demands than the Complex task because it uses fewer stimuli and fewer pairings between stimuli and rules. However,

the Complex task may involve less interference; assigning one rule per pair of stimulus cards (Complex task) probably causes less interference than reversing a rule for the same pair of cards (DCCS). So, although we expected structural similarities between performance on the Complex task and the Advanced DCCS, there were also many reasons to expect performance to differ somewhat.

Experiment 2 further tested whether the rule-based category approach successfully predicts the difficulty of rule use in young children. In this experiment, we examined 5-year-olds' response times in completing the three new rule-use tasks. Response times have been shown to be useful for the study of the categorization of compound stimuli in children, particularly to analyze the complexity of strategies that children may apply (Visser & Raijmakers, 2012). We expected response times across these three tasks to vary with the complexity of the categorization rules that can be abstracted in each of them.

Experiment 1

The main aim of the first experiment was to test whether the rule-based category approach predicts children's performance on rule-use tasks. The Simple, Intermediate, and Complex tasks and the Advanced DCCS task were administered to 124 preschoolers. To prevent children from misunderstanding the instructions, tasks included a training phase for each rule with feedback from the experimenter, followed immediately by a test phase without feedback. To match the procedure across the rule-use tasks and the DCCS, we used the Advanced/Star/Border version of the DCCS (Carlson, 2005; Chevalier & Blaye, 2009; Hongwanishkul et al., 2005). Although the advanced version is more difficult than the standard DCCS (it is still difficult at 5 and 6 years of age), the decision structure on which the rules are based is similar to that of the standard DCCS. The main difference is that the advanced version includes an extra block in which participants need to alternate between the *shape* game and the *color* game. In keeping with the standard DCCS, we used a verbal cue instead of a visual one (e.g., a border) to signal switches between games.

Method

Participants

A total of 124 healthy children (55 boys and 69 girls) were split into three age groups—3-year-olds ($M = 3.5$ years, $SD = 0.26$, $n = 42$), 4-year-olds ($M = 4.4$ years, $SD = 0.30$, $n = 39$), and 5-year-olds ($M = 5.6$ years, $SD = 0.32$, $n = 43$)—from two public schools of the same township (99% of kindergartens are public in France). Most children were from middle-class families. All of the children participated voluntarily, and their parents signed an informed consent form.

Stimuli

For the Simple, Intermediate, and Complex rule-use tasks, there were three sets of cards (randomly associated with the tasks). Each laminated card depicted a bidimensional image (shape and color). The cards in the "Nature" set consisted of two butterflies (one yellow and one green) and two flowers (one yellow and one green). The "Vehicles" set consisted of two motorbikes (one gray and one orange) and two cars (one gray and one orange). The "Cutlery" set consisted of two spoons (one pink and one brown) and two forks (one pink and one brown). The decision to use several stimulus categories (biological kinds, vehicles, etc.) was thought to improve external and construct validity and limit carryover effects. The 10×16 -cm target cards depicted animals and were unrelated to the test cards: a sheep and a cat. For the DCCS task, each target card (a red rabbit and a blue boat) was attached to a box. The laminated test cards (7.0×9.5 cm) depicted blue rabbits or red boats.

Procedure

Each participant performed four tasks (the three rule-use tasks and the DCCS) in random order in a single session.

Rule-use tasks. The following sorting instructions were given once before the three rule-use tasks: “Here we have a sheep and a cat. I am going to tell you which cards the cat likes and which ones the sheep likes. You have to give the cat what he likes and give the sheep what he likes. Do you understand?” If the answer was “yes,” the tasks began. The experimenter repeated that the next task would deal with the cat and the sheep whenever the DCCS was performed between two other tasks. The sheep and cat images were each attached to one of two boxes in which children needed to place the test cards. To minimize the influence of the previous trials, children were asked to place the cards face down when sorting them into the boxes (the face down condition has been shown to be easier than the face-up condition; Kirkham, Cruess, & Diamond, 2003).

Each task was composed of two rule-learning phases followed by a final test phase. In each task, the two rule-learning phases used different cards and featured different rules. In each rule-learning phase, children learned one rule for sorting two kinds of cards (e.g., green flowers, yellow flowers). For example, they were told, “The sheep likes the green flowers and the cat likes the yellow flowers.” Then children sorted a total of 6 cards (three series of the 2 cards such as the green flowers and the yellow flowers) given in random order. Children received feedback in these trials, with the instructions repeated each time an incorrect response was given. After these initial trials, children sorted these same 6 cards again (given in random order) but without feedback from the experimenter. A similar procedure was used for the second rule; children first sorted 6 new cards while receiving feedback (e.g., three series of 2 new cards for which the rule was “The sheep likes the yellow butterflies and the cat likes the green butterflies” if the task was complex) and then sorted a second set of 6 cards without feedback. Finally, in the test phase, children sorted a total of 12 cards (i.e., three series of the 4 cards previously seen given in random order), which required using *both* sets of sorting rules (e.g., “The sheep likes the green flowers and the cat likes the yellow flowers” and “The sheep likes the yellow butterflies and the cat likes the green butterflies”). Fig. 1 recapitulates for each task how the two rule-learning phases succeeded one another, totaling five blocks and 36 cards (6 cards for Rule 1 with feedback, 6 cards for Rule 1 with no feedback, 6 cards for Rule 2 with feedback, 6 cards for Rule 2 with no feedback, and 12 cards for the final test phase in which the two rules were mixed).

Fig. 1 shows the rules children learned and applied in each rule-learning phase and in the final test phase. To illustrate the tasks here, we describe them using the Nature set only. Note, however, that the particular cards used in each task varied across children and across tasks. In all tasks, the first rule-learning phase required children to learn a rule assigning dark flowers to the sheep and white flowers to the cat. The rules in the second rule-learning phase varied across the three tasks. In the Simple task, children learned a parallel rule for the butterflies—again, dark to the sheep and white to the cat. In the Intermediate task, they learned to assign both dark and white butterflies to the sheep, so no objects were assigned to the cat. Finally, in the Complex task, the rules for the butterflies were reversed to those for the flowers—dark butterflies to the cat and white butterflies to the sheep.

Advanced DCCS task. The DCCS was administered similarly to the rule-use tasks, with two rule-learning phases followed by a final test phase. In the first rule-learning phase, children were told rules either for the *same color* game or for the *same object* game (this was determined at random). For example, in the same color game, children were instructed to put blue rabbits into the blue boat box and red boats into the red rabbit box. As in the rule-use tasks, children then completed six trials with feedback from the experimenter, followed by six further trials without feedback. Children then began the second rule-learning phase, which used the rules for the other game (e.g., rules for the *same object* game if the *same color* game was played first). Again, they completed six trials with feedback and a further six trials without feedback. Finally, in the final test phase, children were told, “Now sometimes we are going to play the *same color* game and the other times the *same object* game. You will have to sort the cards by paying attention to which game we are playing. Let’s start.” Before each trial, the experimenter asked, “If we’re playing the *same color* [or *same object*] game, where does this card go [the card was given to children]?” and so forth. Depending on the speed with which children were able to correctly sort the cards, the instructions were sometimes reduced to “We are playing the *color* [or *object*] game.” Given the verbal cue that prompted participants to use the second rule, it was not necessary for the cards to be marked with a visual cue (Carlson, 2005; Chevalier & Blaye, 2009).

Results

An assessment of the normality of data relating to error rates using the Shapiro–Wilk test showed that the number of errors was not normally distributed as a function of age (3, 4, or 5 years) and as a function of phase in each of the tasks. Parametric tests were run in violation of the normality assumption because they are usually very robust against such violations. When appropriate, Gamma and McNemar tests were performed on crosstabs.

A preliminary repeated measures analysis of variance (ANOVA) on mean error rates found no evidence of performance varying across the Vehicle, Nature, and Cutlery stimulus sets, $F(2, 246) = 0.12$, $p = .89$, so this factor was not included in the main analyses. Another preliminary analysis indicated no global order effects. Global order effects were analyzed by recoding the six possible permutations of our three new rule-use tasks by adding 1 to a variable every time a pair of tasks was administered with the purportedly simpler task given first. Using this recoding method simplified the analysis by reducing the six possible permutations of the three tasks to four cases (i.e., scores 0–3). A one-way ANOVA on the global proportion of errors was nonsignificant when this variable was used as the main factor, $F(3, 119) = 1.13$, $p = .34$. When rank effects were analyzed task by task (the factor represented whether the task was performed first, second, third, or last), including the DCCS, none of the four separate one-way ANOVAs was significant either, $F_s(3, 120) < 1.10$, $p > .35$.

We hypothesized that performance in the final test phase, where rules are mixed, would vary among the three new rule-use tasks but not between the Complex task and the Advanced DCCS. Table 1 shows how performance varied across the four tasks. We conducted a 4 (Task: Simple, Intermediate, Complex, or DCCS) \times 3 (Age: 3, 4, or 5 years) repeated measures ANOVA on the proportion of errors in the final test phase, with task as a within-participants factor and age as a between-participants factor. This revealed a significant variation between the mean number of errors in the different tasks, $F(3, 360) = 128$, $p < .001$, $\eta_p^2 = 52\%$. Performance was significantly better in the Simple task ($M = 7\%$, $SD = 1.8$) than in the Intermediate task ($M = 23\%$, $SD = 1.6$), $F(1, 121) = 62.90$, $p < .001$, $\eta_p^2 = 34\%$, and significantly better in the Intermediate task than in the Complex task ($M = 41\%$, $SD = 1.5$), $F(1, 121) = 76$, $p < .001$, $\eta_p^2 = 39\%$. There was no significant difference, however, between the performance on the Complex task and the DCCS ($M = 39\%$, $SD = 1.4$). Age also had an effect on the proportion of errors, $F(1, 120) = 721$, $p < .001$, $\eta_p^2 = 86\%$. Post hoc analyses (Newman–Keuls) showed that 3-year-olds made significantly more errors ($M = 35\%$) than 4-year-olds ($M = 28\%$), who in turn made significantly more errors than 5-year-olds ($M = 18\%$), regardless of task type. We also observed an interaction between age and task type for the proportion of errors, $F(6, 360) = 2.60$, $p = .019$, $\eta_p^2 = 4\%$, which attests to the fact that the gaps between the tasks were smaller for 5-year-olds. When the tasks were analyzed separately

Table 1

Mean percentage of errors observed in the final test phase (fifth block) of the three rule-use tasks (Simple, Intermediate, and Complex) and DCCS in Experiment 1 for the three age groups and mean percentage of errors and correct response times observed in the final test phase (fifth block) of the three rule-use tasks in Experiment 2.

| Experiment 1 | | | | | | | |
|--------------|-------------|----------|----------|-------------|-------|----------|-----------|
| Task | Age (years) | | | ANOVA | | | |
| | 3 | 4 | 5 | $F(2, 121)$ | p | η^2 | Post hoc |
| Simple | 14 (3.0) | 03 (3.1) | 03 (2.9) | 4.6 | <.05 | .07 | 3 < (4,5) |
| Intermediate | 31 (2.7) | 24 (2.8) | 14 (2.7) | 9.2 | <.001 | .13 | (3,4) < 5 |
| Complex | 47 (2.6) | 44 (2.8) | 31 (2.6) | 10.2 | <.001 | .15 | (3,4) < 5 |
| DCCS | 48 (2.4) | 43 (2.5) | 25 (2.4) | 26.5 | <.001 | .31 | (3,4) < 5 |
| Experiment 2 | | | | | | | |
| | % Error | | | Correct RT | | | |
| Simple | 5 (1.4) | | | 1854 (99) | | | |
| Intermediate | 14 (2.9) | | | 1805 (115) | | | |
| Complex | 43 (3.7) | | | 2481 (139) | | | |

Note. Standard errors are given in parentheses.

Table 2
Correlation matrix for Experiment 1.

| | Simple | Intermediate | Complex |
|----------------------------|----------------|----------------|-----------------------|
| <i>Simple</i> | | | |
| DCCS | .267 (.003) | .284 (.001) | .576 (.000) |
| Simple | | .270 (.002) | .225 (.012) |
| Intermediate | | | .263 (.003) |
| <i>Controlling for age</i> | | | |
| DCCS | .180 (.047) | .114 (.210) | .489 (.000) |
| Simple | | .213 (.019) | .157 (.084) |
| Intermediate | | | .154 (.090) |

Note. *p* is given in parentheses and *N* = 124.

to study the effect of age, each one-way ANOVA produced a significant result. One prediction of our approach is that the three complexities can lead to different patterns at each age. The last column of [Table 1](#) shows how the different age groups were differentiated by the post hoc analyses (Bonferroni corrections were made for the pairwise comparisons). The Simple task isolated 3-year-olds, and the other two tasks isolated 5-year-olds. However, children's performance on the Intermediate task might have been overestimated in these analyses because a 75% success rate in this task can be achieved by simply perseverating on the rule from the first rule-learning phase. In the online [Supplementary material](#), we report additional analyses showing the same overall findings when controlling for these concerns.

We also examined whether performance was correlated across the Complex task and the Advanced DCCS. We predicted that performance would be correlated because, according to the rule-based category approach, both tasks depend on the same embedded rule structure. Although we observed slightly better performance on the DCCS than on the Complex task for 5-year-olds (24.61% error rate in the DCCS vs. 31.2% in the Complex task), and this difference was significant, $t(42) = 2.49$, $p = .017$, we obtained a significant relationship based on the number of correct response in each task across participants, $r = .58$, $p < .001$, $N = 124$ (see [Table 2](#)), and this correlation remained significant after controlling for age, $r = .49$, $p < .001$. Although this correlation represents less than 25% of shared variance, it is still the largest correlation observed among the tasks. Although the correlations among the four tasks were all significant (between .225, $p = .012$, and .284, $p = .001$, when controlling for age, the minimal and maximal values were lowered to between .114, $p = .21$, and .213, $p = .019$, and in this case two of them remained significant), we still observed a significant greater correlation between the DCCS and the Complex task than between all of the other pairs. For instance, the *t* using [Steiger's \(1980\)](#) formula was at least $t(123) = 3.46$, $p < .001$, for the correlation between the DCCS and the Complex task and other correlations. This finding supports another prediction of the rule-based category approach.

It is also worth mentioning that the difficulties in the final test phase did not entirely derive from difficulty with switching to the second rule. Effectively, the percentages of errors for the Simple task, the Intermediate task, the Complex task, and the DCCS in the second rule-learning phase were 0.5%, 1.8%, 6.7%, and 6.9%, respectively, across all age groups, which means that the second rule was most often learnable using our procedure. Hence, the difficulty in the final test phase encountered by children cannot be reduced to the slight increase of errors while using the second rule but rather to difficulty in combining the two rules. In addition, another correlation was found between the DCCS and the Complex task, after controlling for age, in the second rule-learning phase when the second rule was tested without any feedback (fourth block), $r = .226$, $p = .012$. This correlation shows that

Table 3

Frequencies of the perseveration, transitional, and switch states in the post-switch phase where children received no feedback for their post-switch behavior for the DCCS and the Complex task.

| DCCS | Complex task | | |
|---------------|---------------|---------------|-----------|
| | Perseverators | Transitionals | Switchers |
| Perseverators | 1 | 0 | 4 |
| Transitionals | 1 | 3 | 3 |
| Switchers | 1 | 8 | 103 |

perseveration on the first rule was similar between the two tasks.¹ The histograms of the number of correct responses by age group during the fourth block for the DCCS and the Complex task are reported in the [Supplementary material](#) and show that most children in all age groups correctly learned the second rule in both tasks. Still, we classified the participants as *switchers* if they obtained five or more correct responses in the fourth block, as *perseverators* if they obtained one correct response or no correct responses, or as *transitional* otherwise. The overall crosstab (Table 3) reporting the frequencies of participants in these different states indicated that among the perseverators of the DCCS, 1, 0, and 4 participants were classified as perseverators, transitional, and switchers, respectively, in the Complex task; the remaining crosstab indicated frequencies equal to 1, 3, and 3 participants, respectively, for the transitional group of the DCCS and 1, 8, and 103 participants, respectively, for switchers of the DCCS. Thus, 107 participants obtained the same classification in both tasks (although this number was inflated by 103 participants who successfully switched in each task), 12 participants differed in only one level of classification, and 5 participants differed in two levels of classification. This crosstab led to a significant test of association ($\text{Gamma} = 2.0, p = .045$). Summing up the crosstab for each task, we observed 3, 11, and 110 children in the respective perseveration/transitional/switch states in the Complex task and 5, 7, and 112 children in the same respective states in the DCCS.

One last analysis aimed at testing whether there were differences in performance on any of the rule-learning phases, especially between the phases that did and did not give feedback, and whether these differences related later to differences in the test phase. Although not always significant, we found a systematic slight decrease in error rates between the feedback phases and the no feedback phases (the maximal difference observed was a 5.5% gain). Among the six paired-sample *t* tests that were run on the mean error rate per block, four tests were significant, $t_s(123) > 2.2, p_s < .05$, but only two of them remained significant after a Bonferroni correction for the family-wise error rate (for the Simple task, between Block 1 and Block 2, and for the Moderate task, between Block 3 and Block 4). When we computed the mean difference between the feedback phases and the no feedback phases for each participant, which we paired to their performance in the final test phase, the correlation was not significant, $r = .02, p = .14, N = 124$.

Discussion

The main predictions of the rule-based category approach were supported. As predicted, performance varied among the three new rule-use tasks; performance was better in the Simple task than in the Intermediate task and better in that task than in the Complex task. We had also predicted that performance would not significantly differ between the Complex task and the advanced version of the DCCS and that performance across these tasks would correlate. These predictions were also supported overall, although with some caveats. When considering all children together, performance did not significantly differ between the Complex task and the DCCS, although for 5-year-olds there was a slight performance advantage for the DCCS. Likewise, performance on these two tasks was correlated. Although the correlation was somewhat weak, it was nonetheless stronger than any other correlation between the tasks administered. As noted earlier, although the Complex task and the DCCS are

¹ Children were judged to perseverate in the post-switch phase of the Complex task if they did not respond correctly. Because the responses needed to be reversed from one rule to the next in this task, children who responded incorrectly applied the first rule instead of the second rule in the post-switch phase.

proposed to depend on the same underlying rule structure, there are many other differences between these tasks (e.g., differences between the stimuli might make the Complex task more taxing on memory); these differences might be responsible for the relatively weak correlation and the performance difference in 5-year-olds. Regardless, the findings provide overall support for the suggestion that the rule-based category approach can predict the relative difficulty of rule-use tasks in young children.

These findings may also be informative about children's difficulty with the DCCS and suggest that this difficulty may principally stem from the complexity of the entire rule system that needs to be kept in mind to sort the cards. We found that performance on the Complex task and the Advanced DCCS was correlated, which is consistent with the possibility that both tasks depend on similar underlying rule structures. As such, the findings cast doubt on certain explanations for children's difficulty with the DCCS. For example, one explanation for poor performance on the DCCS is that young children have difficulty in redescribing the bivalent cards in multiple ways according to the different rules in the DCCS (Kloo & Perner, 2003, 2005; Perner & Lang, 2002; Zelazo et al., 2003b). However, our Complex task does not make these demands (i.e., each card leads to one category unequivocally and without the need for redescription) but is as difficult as the Advanced DCCS, and we found that performance on the tasks is correlated. This suggests that performance on the DCCS might not be strongly related to the ability to describe stimuli in multiple ways. Another explanation for difficulty with the DCCS is that young children are unable to disengage from the pre-switch rule, which they continue to apply during the post-switch phase in the DCCS. Indeed, it can be difficult to disengage from the first rule to load the elements of the second rule into working memory (Bialystok & Martin, 2004; Hanania, 2010; Kirkham et al., 2003; in adults, see Allport, Styles, & Hsieh, 1994; Allport & Wylie, 2000; Meiran, 1996; Monsell, 2003). However, our Intermediate task also required children to disengage from a pre-switch rule, and performance on this task was superior to performance on the Complex task or DCCS. In addition, performance was particularly strong on both the Complex task and the DCCS when the second rule was tested without any feedback (fourth block).² These results suggest that perseveration on the first rule cannot entirely account for children's difficulty in coordinating the two rules in the final test phases of these tasks.

However, accounts of DCCS performance that highlight the role of working memory (Cepeda & Munakata, 2007; Chevalier & Blaye, 2008; Morton & Munakata, 2002) might better account for this result because in the Intermediate task the less complicated sub-rule is easier to load. These claims are all based on the conclusion that the observed similarities between performance on the Complex task and the Advanced DCCS result because the tasks are represented similarly and may draw on the same (or overlapping) psychological abilities; in the General Discussion, we acknowledge the need for caution in concluding this.

Experiment 2

The objective of the second experiment was to replicate the ordering of the three rule-use tasks (Simple, Intermediate, and Complex) found in the first experiment but with a refined analysis of performance based on response times (RTs). RT is a particularly interesting dependent variable to show that the rules instructed in the current study do not merely reduce to rote learning. If

² The percentages of errors for both the Complex task and the DCCS were less than 7% in this particular phase. This is less than in previous studies that used no feedback but is still consistent with a recent study showing that feedback on post-switch behavior is helpful even in 3-year-olds (van Bers, Visser, & Raijmakers, 2014b). Children are usually not provided with corrective feedback during the post-switch trials, and it is only children who pass the post-switch phase of the standard version of the DCCS who can proceed to the advanced version (Zelazo, 2006). In our version, all children were provided with corrective feedback during the first six trials of the post-switch phase, which might explain why most of the children correctly classified the post-switch cards after this phase. The number of errors was a bit higher during the first six trials with feedback, amounting to a global 14.5% of errors (23% in 3-year-olds) instead of 6.9% for the next six trials without feedback. Performance on the first six trials is still quite low, but again this can be explained by the feedback that was given throughout the phase. In 3-year-olds, 27 children obtained five or six correct responses in the post-switch phase that included feedback, representing 64% of the children, which is comparable to the 85% (instead of 38% in the control condition) obtained by van Bers and colleagues (2014b) in their condition with feedback. Altogether, this result suggests that perseveration on the first rule cannot entirely account for children's difficulty in coordinating the two rules in the final test phases of these tasks.

participants simply engaged in rote learning, we would obtain similar RTs for all of the to-be-categorized objects (i.e., all of the objects across rules and tasks would be classified using one branch per object). Instead, we expected that RTs would vary across the three tasks and that they would be predicted by the complexity of the rule structures that can be abstracted from the different task instructions (i.e., because more complex rules require a greater number of decision steps) struc., because more complex rule dictated for the asks ficulty, ting of Experiment 1 results already. nistered. Complex task. ari. Another prediction based on RTs is related to the effect of rule complexity on mixing costs. If rule complexity is the complexity of the rule structures that can be abstracted from the different task instructions, mean difference in RTs between using a single rule and using mixed rules should increase with rule complexity because more complex rules are based on more conflicting sub-rules (i.e., on average, the number of features that are different from one sub-rule to another increases along the Simple, Intermediate, and Complex task dimension). Because collecting RT data was our chief interest in this experiment, we tested children who would find the tasks to be quite feasible and who would pass on most trials. For this reason, we tested a group of 5- and 6-year-olds.

Method

Participants

A total of 36 preschool children ($M_{\text{age}} = 5.7$ years, $SD = 0.4$, 15 boys and 21 girls different from those in Experiment 1) from two different classes of two different public schools voluntarily participated in this study. Most children were from middle-class families. Consent forms were distributed to all parents, and we included only participants for whom we had received parental consent.

Materials and stimuli

Computerized versions of the tasks were administered on a laptop computer. The categorization tasks were developed and run using E-Prime 2 software (Psychology Software Tools).

Three different sets of four stimuli were randomly used in the three tasks, with the constraint that each task was associated with the sets in equal proportion. The three sets of four compound stimuli measuring 5×5 cm were constructed using a combination of two specific shapes and two specific colors in each set. The first set, consisting of simple images (blue square, red square, blue circle, and red circle), was used to illustrate the procedure; the other two sets were based on stars, pentagons, and two abstract shapes from the Jessica S. Horst database (<http://www.sussex.ac.uk/wordlab/noun>), and different shadings and patterns were used instead of simple colors (e.g., horizontal or vertical lines, shades of gray). These three sets were simply built to induce minimal interference regarding features.

Procedure

Children completed the Simple, Intermediate, and Complex tasks, with task order randomized across participants. During testing, children sat facing the laptop computer. The experimenter sat close to children to give the instructions during the three tasks. Children were instructed that they would give pictures to either Mickey Mouse or Donald Duck by pressing the “Mickey” or “Donald” keys on the keyboard. On each key was a colored sticker showing the face of the corresponding character. Children were told to use both hands during the tasks and to press the keys with their index fingers. They were also instructed to do their best to avoid errors.

The computerized versions of the tasks differed somewhat from the versions in Experiment 1. In the current versions, children first learned and practiced categorization rules in four blocks, each with 8 trials, and then completed a final test block with 16 trials. Before each block, instructions were given verbally by the experimenter with a corresponding image on the computer screen. In explaining the instructions, the experimenter referred to stimuli without naming or describing them (e.g., “This picture goes to Mickey,” “These pictures go to Donald”). In the corresponding image (shown while the instructions were explained), the pictures were organized in rows. Each row was followed by an arrow pointing to a picture of Mickey or Donald to show the category to which each object was assigned. The instruction before the final test block reminded participants of both rules simultaneously.

Although some aspects of the procedure differed from Experiment 1, the rule structures of the tasks remained the same. In Block 1 children learned a rule for categorizing two stimuli, and in Block 2 they learned the other rule used for categorizing the other two stimuli. In these blocks, children received feedback after each trial; after correct responses the word “OK” appeared in green at the center of the screen, and after incorrect responses the word “FALSE” appeared in red. These words appeared for 1500 ms, and the experimenter also read the feedback to participants. In the remaining blocks, no feedback was given. In Block 3 children applied the first rule again, and in Block 4 they applied the second rule. Finally, children completed the final test block, which required them to categorize all four kinds of objects and to switch between the two rules.

Within each of the five blocks, the stimuli were presented sequentially and randomly, although each stimulus appeared the same number of times (i.e., in Blocks 1–4, each of the two stimuli appeared four times per block; in the final test block, each of the four stimuli appeared four times). A fixation cross appeared and remained in the middle of the screen for 1000 ms before each stimulus appeared. The total length of the experiment was approximately 20 min, including instructions and breaks.

Results

We conducted a series of repeated measures ANOVAs with task (Simple, Intermediate, or Complex) as the main independent variable and error rates and RTs as the two dependent variables. We also used mixing cost (i.e., the mean difference in RTs between using a single rule and using mixed rules) as another dependent variable.

Error rates

A preliminary repeated measures ANOVA on mean error rates found no evidence of performance varying across the three stimulus sets, $F(2, 70) = 2.09$, $p = .13$, so this factor was not included in the main analyses. As in Experiment 1, global order effects were analyzed by recoding the six possible permutations of our three rule-use tasks into a permutation variable with scores ranging from 0 to 3. A one-way ANOVA on the global proportion of errors across the rule-use tasks was nonsignificant when permutation was used as the main factor, $F(3, 32) = 0.50$, $p = .68$. When rank effects were analyzed task by task (the factor represented whether the task was performed first, second, third, or last), none of the four separate one-way ANOVAs was significant either, $F_s(2, 33) < 2.40$, $p > .10$.

We expected error rates to vary across the three tasks (Table 1). A repeated measures 1×3 ANOVA, with task complexity as a within-participants variable, revealed significant variation between the mean number of errors in the final test blocks of the tasks, $F(2, 70) = 56.60$, $p < .001$, $\eta_p^2 = 62\%$. Pairwise comparisons based on the Bonferroni correction showed that the mean percentage of errors in the Simple task ($M = 5\%$, $SD = 1.4$) was significantly lower ($p = .003$) than in the Intermediate task ($M = 14\%$, $SD = 2.9$), which in turn was significantly lower ($p < .001$) than in the Complex task ($M = 43\%$, $SD = 3.7$).

As in Experiment 1, these analyses may overestimate children's performance in the Intermediate task because use of a global univalent rule (i.e., perseverating by the first rule) would allow children to respond correctly on 75% of the trials. As in Experiment 1, we report additional analyses showing the same overall findings when controlling for these concerns in the [Supplementary material](#).

Similar to Experiment 1, a further analysis aimed at testing whether there were differences in performance between the phases that did not give feedback and the phases that gave feedback and whether these differences related later to differences in the test phase. We found a single significant 12% gain for the Complex task between Block 1 and Block 3, $t(35) = 2.30$, $p = .026$, but this result did not hold significant after a Bonferroni correction for the family-wise error rate. In addition, as in Experiment 1, when we computed the mean difference between the feedback phases and the no feedback phases for each participant, which we paired to their performance in the final test phase, the correlation was not significant, $r = .08$, $p = .65$, $N = 36$.

Correct response times

We examined RTs only for trials in which correct responses were given. Overall, RTs in the final block varied from 1854 ms ($SD = 596$) for the Simple task, to 1805 ms ($SD = 689$) for the Intermediate

task, to 2481 ms ($SD = 824$) for the Complex task, $F(2,68) = 18.40$, $p < .001$, $\eta_p^2 = .35$.³ Although we expected significant differences between each pair of tasks, post hoc comparisons using Bonferroni correction did not show a significant difference between the Simple task and the Intermediate task. Nonetheless, RTs were significantly greater in the Complex task than in both the Simple task ($p < .001$) and the Intermediate task ($p < .001$). However, opportunities for getting faster RTs are greater in the Intermediate task because three cards belong to the same category. In the [Supplementary material](#), we report additional analyses that better differentiate the Intermediate task from the Simple and Complex tasks when trials were selected to match the decision structures.

Mixing cost

Although many components can be computed in task switching, including alternation costs, mixed list costs, and switch costs (Meiran, 2000), we chose to compute a simple mixing cost for every participant and each task; this was calculated with the formula $RT_Block5 - (RT_Block3 + RT_Block4)/2$. The repeated measures ANOVA showed that the mixing costs (-31 ms for the Simple task, 382 ms for the Intermediate task, and 509 ms for the Complex task) increased significantly with task difficulty, $F(2,58) = 4.90$, $p = .01$, $\eta_p^2 = .15$. Using the Bonferroni adjustment, the post hoc comparisons showed a single significant difference between the Simple task and the Complex task ($p = .04$), although we also obtained a significant difference between the Simple task and the Intermediate task ($p = .02$) using Fisher's LSD (least significant difference).

Latent Markov modeling

In this experiment, we also collected trial-by-trial data. One advantage of trial-by-trial data is that they allow an analysis of accuracy in the post-switch phase that is based on latent Markov models. This method is more powerful than simply reporting frequencies based on some arbitrary pass/fail criterion.⁴ According to van Bers and colleagues (2011), there are three main categories of participants usually found for the DCCS, namely that (a) participants switch correctly and make (almost) all items correct, (b) they perseverate, or (c) they show a transition from perseveration to switching. Because the current experiment focused on 5- and 6-year-olds, we did not expect to find children who predominantly perseverate. We expected most children to switch correctly and a small transitional group to be shifting from perseverating to switching.

Following van Bers and colleagues (2011), we fit different latent Markov models to the trial-by-trial accuracy data of the Complex task in the post-switch phase. We used the `depmixS4` package (Visser & Speekenbrink, 2010) for the R statistical programming environment (R Development Core Team, 2009) and based our analysis on Block 4, in which no feedback was provided to children. As in Experiment 1, a first simple analysis of the Complex task sought to classify children as perseverators (accuracy ≤ 1 correct response across the entire post-switch block), as in transition (accuracy between 2 and 6 correct responses), or as switchers (accuracy on at least 7 of 8 responses; these children were considered to have passed the post-switch phase). Of the 36 children, 3 were classified as perseverators, 6 were classified as in transition, and 27 were classified as switchers. We then split the 36 children into two groups around the median age (70 months). However, the crosstab reporting the frequencies of participants in the different states according to age group did not lead to a significant test of association between the two ordinal variables ($\Gamma = -.16$, *ns*). Splitting children showing each response pattern by age: perseverators included 1 younger child and 2 older children, children in transition included 2 younger children and 4 older ones, and switchers included 15 younger children and 12 older ones.

Transitions from one response to the next were coded as “s” (switch) when correct, “p” (perseveration) when incorrect, “f” (forward) when correct but preceded by an incorrect response, or “b” (backward) when incorrect but preceded by a correct response. (Although the extra “f” and “b”

³ The degrees of freedom is 68 instead of 70 because one of the participants obtained absolutely no correct response in the Complex task in the final block, although this participant did pass the previous blocks; one simple explanation is that the participant systematically reversed the two rules.

⁴ We did not take this approach in Experiment 1 because in that experiment we neglected to record children's responses trial by trial.

codes are not necessary to study the patterns of responses, they highlight the directions of transitions in responses.) Overall, we found only 12 different patterns, and these allowed us to categorize children into three groups: 27 switchers (24 ssssssss, 1 ssssssb, 1 ssssfss, and 1 ssfbssss), 6 transitional (1 pfsbbb, 1 sssbbbbb, 2 sssbbb, 1 sssbbbf, and 1 sbbbbb), and 3 perseverators (1 pbbbbbbb, 1 pbbbbbf, and 1 sbbbbb). Across individuals, there were overall 10 chances out of 213 to observe a backward switch and 5 chances out of 39 to observe a forward switch. The 2.6 odds ratio computed on these numbers shows that it was 2.7 times more likely to observe a forward switch than a backward switch, but this estimation is made globally across individuals and does not take into account individual estimations.

More precise estimates of the transitional probabilities based on these 36 individual patterns were obtained by running an analysis based on latent Markov models. This analysis was inspired by van Bers and colleagues (2011). Although simpler than in the original study, the analysis suggests that the best model for the current data is a two-state model (including a perseveration state and a switch state) in which the probability of a shift from the perseveration state to the switch state was .08 instead of .04 for the reverse switch. Our probability (.08) is lower than the one found by van Bers and colleagues (.15), probably because our children were older (children were 52 months old in the van Bers et al.'s study). The initial probability of a switch state was .94 (instead of .98 in van Bers et al.), and the initial probability of a perseveration state was .06 (instead of .01 in van Bers et al.). The probability to keep perseverating was .92, and the probability to keep switching was .96. The log likelihood (-64.7 , $df = 5$) and two information criteria (AIC [Akaike] = 139.4; BIC [Bayesian] = 157.7) were lower than in the model where the transition probabilities between different states were fixed to zero (log likelihood = -83.6 , $df = 3$; AIC = 177.3; BIC = 195.6) and lower than in a one-state model (log likelihood = -128 , $df = 1$; AIC = 258; BIC = 262). When testing the two closest models, we found a significant chi-square difference between the restricted model (one state) and the more general model (two states with transitional probabilities allowed), $\chi^2(5-1) = -2 [-83.6 - (-64.7)] = 37.8$, $p < .001$.

Discussion

As in Experiment 1, error rates varied among the three rule-use tasks; performance was better on the Simple task than on the Intermediate task and better on that task than on the Complex task. When examining RTs, we also observed differences among the tasks, although the main analysis found significant differences only between the Complex task and the Simple and Intermediate tasks. We also found that most of the children tested in this experiment successfully switched to the second rule in the post-switch phase, which again suggests that perseveration on the first rule cannot entirely account for children's difficulty in coordinating the two rules in the final test phases of the Complex task.

General discussion

The findings of these two experiments provide evidence that the rule-based category learning approach (e.g., Minda et al., 2008) can be used to successfully predict rule use in preschool-aged children. In a first experiment, we examined children's performance on three new rule-use tasks. In each of the three new tasks, children were given instructions for sorting four kinds of stimuli. The instructions were matched in the number of to-be-classified cards, so it might be expected that children should perform similarly on all three new tasks. Instead, performance varied depending on the complexity of the rule structure that could be abstracted from the instructions. In this experiment, children also completed the Advanced DCCS task, a version of the DCCS in which children are randomly cued to switch between the two rules on multiple trials instead of needing to switch between the rules only once (e.g., Carlson, 2005). Although this task featured quite different instructions from those used in our Complex task, we found that performance on these two tasks was correlated and mostly comparable. This was predicted by the rule-based category approach because according to that approach both tasks depend on the same underlying rule structure.

A second experiment replicated the finding that children's performance on the three new rule-use tasks is predicted by the complexity of the rules that can be abstracted in each task. Moreover, this

experiment showed that response times also vary with this complexity, although we found no systematic significant result isolating the three tasks according to their complexity. These latter findings are broadly consistent with previous studies suggesting that task complexity predicts response times (e.g., Deák, Ray, & Pick, 2004; Emerson & Miyake, 2003; Rubinstein, Meyer, & Evans, 2001).

Together, these findings bolster claims that at least some of children's difficulty in rule use derives from limits of their ability to represent rule structures (Zelazo & Frye, 1998). The findings likewise suggest that children do not passively follow instructions in rule-use tasks; if children had simply followed the task instructions by rote, they should have performed similarly in our Simple, Intermediate, and Complex tasks because the instructions in all three tasks specified how to categorize four kinds of stimuli. Instead, children's performance was predicted by the complexity of the rule structures that can be abstracted from these instructions, which implies that children do abstract or generalize a rule structure from instructions. These findings are in line with other studies showing that children, like human adults, can better transcend association-based processes to form abstractions (e.g., Shepard et al., 1961). These findings also connect to the idea that young children who perform well on our tasks can form solid abstract representations that are not stimulus specific (otherwise, there would be no minimization of the rules and children would associate each of the stimuli to only one category by rote learning) and that these representations might be used to generalize to novel stimuli (Kharitonova, Chien, Colunga, & Munakata, 2009; Kharitonova & Munakata, 2011; van Bers, Visser, & Raijmakers, 2014a).

These findings are also important for several other reasons. First, they provide preliminary evidence that children's rule use depends on processes also involved in rule-based category learning. As such, the findings suggest a connection between two different aspects of cognition (i.e., rule use and rule-based category learning) that have been studied extensively but separately from one another. One limitation of the rule-based category learning field has been that categorization was generally studied through classification rather than category use (Markman & Ross, 2003), although a few more recent models can better connect both aspects by integrating category learning and inference mechanisms (e.g., Love, Medin, & Gureckis, 2004). In most studies of rule-based categorization, adult participants attempted to learn complex rule-based categories, varying in three perceptual dimensions, and participants were not usually tested after the learning criterion had been reached (e.g., Feldman, 2000; Nosofsky et al., 1994; Shepard et al., 1961). We used simpler rule structures in investigating children, and the children were explicitly told the rules they needed to apply. We found that children's rule use was predicted by the complexity of the rule structures that could be abstracted from the instructions; young children were successful in using simpler rule structures and produced more errors when using more complex rules.

Second, as noted above, the findings may be informative about children's difficulty with the DCCS. Specifically, the findings are consistent with the possibility that difficulty on the task principally stems from the complexity of the rule system that needs to be represented. In our first experiment, performance on the Complex task and the Advanced DCCS was correlated, which was predicted because both tasks are posited to depend on similar underlying rule structures. As discussed earlier, these findings might be difficult to explain given other accounts of children's difficulty with the DCCS, including accounts claiming that poor performance results from children having difficulty in redescribing bivalent cards according to different rules (Kloo & Perner, 2003, 2005; Perner & Lang, 2002; Zelazo et al., 2003b) and accounts claiming that difficulty results from an inability to disengage from the pre-switch rule. However, drawing these connections assumes that the observed similarities between performance on our Complex task and the Advanced DCCS arise because both tasks draw on the same representational resources, and although this possibility is consistent with the current findings, the findings stop short of showing this and other explanations for the similar performance are possible.

Third, the findings suggest that the development of rule use may be quite protracted. With age, children became more successful in rule-use tasks depending on increasingly complicated rule structures. Although our experiments examined children's ability to apply only three rule structures, many more complicated rule structures could be used to measure rule use in older children and in adults (e.g., Feldman, 2003a). Aside from the three rule structures used in the current study, 13 rule structures are possible with three-dimensional stimuli and more than 200 rule structures are possible with four-dimensional stimuli. The complexity of each of these rule structures can be indexed (as discussed

in the Introduction) and used to study the development of rule use in older children and adults and to refine the study of switch costs that have been limited to a single rule structure so far (Diamond & Kirkham, 2005). Based on the current findings, it might be conjectured that complexity will predict both the difficulty of rule use and the difficulty of switching with these other more complicated rule structures. Answering these questions will be an exciting project for future research.

Acknowledgments

This research was supported in part by a grant from the Agence Nationale de la Recherche (Grant ANR-09-JCJC-0131-01) awarded to Fabien Mathy and by a grant from the Région de Franche-Comté (AAP2010) awarded to Fabien Mathy and Lucie Laurent with the support of the Maison des Sciences de l'Homme et de l'Environnement Ledoux (USR 3124). Ori Friedman's contributions to this research were funded by a grant from the Natural Sciences and Engineering Research Council of Canada.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jecp.2014.10.008>.

References

- Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 421–452). Cambridge, MA: MIT Press.
- Allport, A., & Wylie, G. (2000). Task-switching: Stimulus–response bindings and negative priming. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII: Control of cognitive processes* (pp. 35–70). Cambridge, MA: MIT Press.
- Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the Dimensional Change Card Sort task. *Developmental Science*, 7, 325–339.
- Bourne, L. E. J. (1970). Knowing and using concepts. *Psychological Review*, 77, 546–556.
- Bradmetz, J. I., & Mathy, F. (2008). Response times seen as decompression times in Boolean concept use. *Psychological Research/ Psychologische Forschung*, 72, 211–234.
- Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York: John Wiley.
- Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, 28, 595–616.
- Cepeda, N. J., & Munakata, Y. (2007). Why do children perseverate when they seem to know better: Graded working memory, or directed inhibition? *Psychonomic Bulletin and Review*, 14, 1058–1065.
- Chevalier, N., & Blaye, A. (2008). Cognitive flexibility in preschoolers: The role of representation activation and maintenance. *Developmental Science*, 11, 339–353.
- Chevalier, N., & Blaye, A. (2009). Setting goals to switch between tasks: Effect of cue transparency on children's cognitive flexibility. *Developmental Psychology*, 45, 782–797.
- Deák, G. O., Ray, S. D., & Pick, A. D. (2004). Effects of age, reminders, and task difficulty on young children's rule-switching flexibility. *Cognitive Development*, 19, 385–400.
- Diamond, A., & Kirkham, N. (2005). Not quite as grown-up as we like to think. *Psychological Science*, 16, 291–297.
- Diamond, A., Kirkham, N., & Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Developmental Psychology*, 38, 352–362.
- Emerson, M. J., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language*, 48, 148–168.
- Feldman, J. (2000). Minimization of Boolean complexity in human concept learning. *Nature*, 407, 630–633.
- Feldman, J. (2003a). A catalog of Boolean concepts. *Journal of Mathematical Psychology*, 47, 75–89.
- Feldman, J. (2003b). The simplicity principle in human concept learning. *Current Directions in Psychological Science*, 12, 227–232.
- Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, 10, 483–527.
- Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3½ to 7 years old on a Stroop-like day–night test. *Cognition*, 53, 129–153.
- Hanania, R. (2010). Two types of perseveration in the Dimension Change Card Sort task. *Journal of Experimental Child Psychology*, 107, 325–336.
- Hanania, R., & Smith, L. B. (2009). Selective attention and attention switching: Towards a unified developmental approach. *Developmental Science*, 13, 622–635.
- Hongwanishkul, D., Happaney, K. R., Lee, W. S. C., & Zelazo, P. D. (2005). Assessment of hot and cool executive function in young children: Age-related changes and individual differences. *Developmental Neuropsychology*, 28, 617–644.
- Hovland, C. (1966). A communication analysis of concept learning. *Psychological Review*, 59, 461–472.
- Kharitonova, M., Chien, S., Colunga, E., & Munakata, Y. (2009). More than a matter of getting “unstuck”: Flexible thinkers use more abstract representations than perseverators. *Developmental Science*, 12, 662–669.
- Kharitonova, M., & Munakata, Y. (2011). The role of representations in executive function: Investigating a developmental link between flexibility and abstraction. *Frontiers in Psychology*, 2, 347. <http://dx.doi.org/10.3389/fpsyg.2011.00347>.

- Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimensional switching task. *Developmental Science*, 6, 449–467.
- Kloo, D., & Perner, J. (2003). Training transfer between card sorting and false belief understanding: Helping children apply conflicting descriptions. *Child Development*, 74, 1823–1839.
- Kloo, D., & Perner, J. (2005). Disentangling dimensions in the dimensional change card-sorting task. *Developmental Science*, 8, 44–56.
- Lafond, D., Lacouture, Y., & Mineau, G. (2007). Complexity minimization in rule-based category learning: Revising the catalog of Boolean concepts and evidence for non-minimal rules. *Journal of Mathematical Psychology*, 51, 57–74.
- Levine, M. (1966). Hypothesis behavior by humans during discrimination learning. *Journal of Experimental Psychology*, 71, 331–338.
- Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review*, 111, 309–332.
- Markman, A. B., & Ross, B. H. (2003). Category use and category learning. *Psychological Bulletin*, 129, 592–613.
- Mathy, F. (2012). La compression de l'information dans l'apprentissage de catégories artificielles. *L'année Psychologique*, 112, 593–629.
- Mathy, F., & Bradmetz, J. (2004). A theory of the graceful complexification of concepts and their learnability. *Current Psychology of Cognition*, 22, 41–82.
- Mathy, F., Haladjian, H. H., Laurent, E., & Goldstone, R. L. (2013). Similarity–dissimilarity competition in disjunctive classification tasks. *Frontiers in Psychology*, 4, 26. <http://dx.doi.org/10.3389/fpsyg.2013.00026>.
- Medin, D. L., & Schaffer, M. (1978). A context theory of classification learning. *Psychological Review*, 85, 207–238.
- Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1423–1442.
- Meiran, N. (2000). Modeling cognitive control in task switching. *Psychological Research Psychologische Forschung*, 63, 234–249.
- Minda, J., Desroches, A. S., & Church, B. A. (2008). Learning rule-described and non-rule-described categories: A comparison of children and adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1518–1533.
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7, 134–140.
- Morton, J. B., & Munakata, Y. (2002). Are you listening? Exploring a developmental knowledge action dissociation in a speech interpretation task. *Developmental Science*, 5, 435–440.
- Nosofsky, R. M., Gluck, M. A., Palmeri, T. J., McKinley, S. C., & Gauthier, P. (1994). Comparing models of rules-based classification learning: A replication and extension of Shepard, Hovland, and Jenkins (1961). *Memory & Cognition*, 22, 352–369.
- Perner, J., & Lang, B. (2002). What causes 3-year-olds' difficulty on the Dimensional Change Card Sorting task? *Infant and Child Development*, 11, 93–105.
- R Development Core Team (2009). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. URL: <http://www.R-project.org/>.
- Rehder, B., & Hoffman, A. B. (2005). Thirty-something categorization results explained: Selective attention, eyetracking, and models of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 811–829.
- Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 763–797.
- Shepard, R. N., Hovland, C. L., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs*, 75(13) [Whole No. 517].
- Smith, J. D., Minda, J. P., & Washburn, D. A. (2004). Category learning in rhesus monkeys: A study of the Shepard, Hovland, and Jenkins (1961) tasks. *Journal of Experimental Psychology: General*, 133, 398–414.
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87, 245–251.
- van Bers, B. M. C. W., Visser, I., & Raijmakers, M. (2014a). Preschoolers can form abstract rule representations regardless of cognitive flexibility. *Journal of Experimental Child Psychology*, 124, 50–66.
- van Bers, B. M. C. W., Visser, I., & Raijmakers, M. (2014b). Preschoolers learn to switch with causally related feedback. *Journal of Experimental Child Psychology*, 126, 91–102.
- van Bers, B. M. C. W., Visser, I., van Schijndel, T. J. P., Mandell, D. J., & Raijmakers, M. E. J. (2011). The dynamics of development on the Dimensional Change Card Sorting task. *Developmental Science*, 14, 960–971.
- Vigo, R. (2006). A note on the complexity of Boolean concepts. *Journal of Mathematical Psychology*, 50, 501–510.
- Visser, I., & Raijmakers, M. E. J. (2012). Developing representations of compound stimuli. *Frontiers in Developmental Psychology*, 3, 73. <http://dx.doi.org/10.3389/fpsyg.2012.00073>.
- Visser, I., & Speekenbrink, M. (2010). DepmixS4: An R-package for hidden Markov models. *Journal of Statistical Software*, 36, 1–21 [Current version available from CRAN; URL: <http://www.jstatsoft.org/v36/i07/>].
- Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, 1, 297–301.
- Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives*, 6, 354–360.
- Zelazo, P. D., & Frye, D. (1998). Cognitive complexity and control: II. The development of executive function in childhood. *Current Directions in Psychological Science*, 7, 121–126.
- Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development*, 11, 37–63.
- Zelazo, P. D., Müller, U., Frye, D., & Marcovitch, S. (2003a). The development of executive function in early childhood: I. The development of executive function. *Monographs of the Society for Research in Child Development*, 68(3) [Serial No. 274].
- Zelazo, P. D., Müller, U., Frye, D., & Marcovitch, S. (2003b). The development of executive function in early childhood: III. Study 2: Rule complexity and stimulus characteristics in executive function. *Monographs of the Society for Research in Child Development*, 68(3), 48–64 [Serial No. 274].